

REMARKS

Applicants gratefully acknowledge Examiner Ngo for taking time on February 1, 2007, to participate in a telephone interview including co-inventor Gustavson, who presented a summary of the present invention and its distinction from the cited reference Myzewski (US Patent No. 5,099,447). The Examiner indicated that he would consider points brought up by the presentation in this next evaluation and Applicant invited the Examiner to call Applicants' representative if he had additional technical questions during the course of the evaluation so that subsequent interviews could be arranged to address the Examiner's questions.

Entry of this Amendment is proper under 37 CFR §1.116, since no new claims or new issues are raised relative to the rejection based on Myzewski and the claim revisions of the independent claims attempt to find wording that are more acceptable to this Examiner relative to the statutory subject matter rejection, thereby reducing the issues for Appeal.

Claims 1, 2, and 4-27 are all the claims presently pending in the application. Claim 3 is canceled above.

It is noted that composite blocking is the more generic concept of double blocking addressed in claims 4, 11, 17, 20, 24, and 27, which claims are not rejected under the prior art evaluation currently of record. Applicants, therefore, conclude that the Examiner considers these claims to be allowable pending resolution of the issue of non-statutory subject matter.

It is noted that Applicants specifically state that no amendment to any claim herein should be construed as a disclaimer of any interest in or right to an equivalent of any element or feature of the amended claim.

Claims 1-27 stand rejected under 35 U.S.C. § 101 as allegedly directed to non-statutory subject matter. Claims 1-3, 6, 7, 9, 10, 12, 14, 15, 18, 23, 25, and 26 stand rejected under 35 U.S.C. § 102(b) as allegedly anticipated by U.S. Patent No. 5,099,447 to Myszewski. Claims 5, 8, 13, 16, 19, 21, and 22 stand rejected under 35 U.S.C. § 103(a) as allegedly unpatentable over Myszewski.

These rejections are respectfully traversed in the following discussion.

I. THE CLAIMED INVENTION

The claimed invention is directed to method of increasing at least one of efficiency and speed in executing a matrix subroutine on a computer. Data for a matrix subroutine call is stored as contiguous data in a computer memory in an increment block size that is based on a cache size of the computer. A first dimension of the block is larger than a corresponding first dimension of the cache and a second dimension of the block is smaller than a corresponding second dimension of the cache, so that the block fits into a working space of the cache.

Since standard Fortran and C two-dimensional arrays are stored in memory/processed in either a column or a row of the matrix data, depending on which language is used, there are inefficiencies in retrieving data for sub-blocks of matrix data, as discussed in more detail below.

In contrast, the present invention teaches that the data is to be stored in memory as a collection of sub-blocks of data that each fits into the cache working area and that is stored contiguously in memory, to preclude thrashing of data in cache.

By actually storing in memory the matrix data as contiguous submatrix data and by sizing the submatrix block of data so that one dimension of the submatrix block exceeds the dimension of the cache relative to the size of the cache (a number of submatrix blocks make up a square matrix), the processing is executed faster than conventional methods, since data thrashing is reduced or eliminated.

Another aspect of the present invention is the recognition that many engineering/scientific matrices are typically directed toward square matrices, since such square matrices express linear equations in which a specific solution can be found. In contrast to conventional wisdom, the present invention also teaches that such square matrices can be viewed as rectangular blocks of matrix data that can be brought into a working area of cache as a block of contiguous data, again, thereby reducing or eliminating the thrashing that typically occurs when the entire square block is retrieved. The composite block of the present invention allows matrix data that, in modern engineering/scientific applications, typically involve a matrix having at least one dimension that exceeds either dimension of the working area of the cache, to be brought into cache in contiguous blocks that address the problem that plagues the art of cache data thrashing.

As an example of increased processing efficiency, an application related to the square blocking of the present invention using contiguous DGEMM operands is Cholesky factorization.

Cholesky factorization is an example of a DLAAFA. On the IBM Power3 processor the implementation of Cholesky factorization using the present invention achieves 92% of peak performance whereas conventional full format LAPACK DPOTRF achieves 77% of peak performance. Moreover, all programming for the new data structures discussed in the present invention can be accomplished in standard Fortran (or C/C++), through the use of higher dimensional full format arrays. Thus, no new compiler support is necessary to implement the present invention.

II. THE 35 USC §101 REJECTIONS

Claims 1-27 stand rejected under 35 U.S.C. §101 as allegedly directed toward non-statutory subject matter. More specifically, the Examiner considers that claims 1-9 and 19-22 are *"... directed to a computer implemented method of calculation where the inputs are numbers and the results are numbers."* Claims 10-14 are although directed to an apparatus broadly encompass a general computer implementing the method. In order for a claimed invention that is directed to such a computer implemented method of calculation, or an apparatus that is no more than a general computer implementing the method of calculation to be statutory, the claimed invention must accomplish a practical application. That is the claimed invention must transform an article or physical object to a different state or thing, or produce a useful, concrete and tangible result Claims 15-18 are rejected under 35 U.S.C. 101 as being directed to signal bearing medium including carrier signal which is non-statutory subject matter. It should be noted that claims 15-18 if merely amended to limit the invention to a computer readable medium that excludes carrier wave or any form of propagation signals would also be rejected under 35 U.S.C. 101 for the same reason set forth in paragraph 3 above."

In Paragraph 7 on page 4 of the Office Action, the Examiner responds to Applicants' arguments in the previous Amendment, as follows:

"First, it is respectfully submitted that the final result produced by the invention as recited in the claims is merely numerical values, and thus does not constitute a useful result as in State St. Bank to Trust Co., 149 F.3d at 1373, where the result indicate dollar amounts to a final share price. Further, it should be noted that in determining whether the invention produces a useful, tangible, and concrete result, the focus is not on whether the steps taken to achieve a

particular result are useful, tangible, and concrete, but rather the final result achieved by the claimed invention is “useful, tangible, and concrete”, see “Interim Guidelines for Examination of Patent Applications for Patent Subject Matter Eligibility”, MPEP 2106. Therefore, the claimed invention is directed to non-statutory subject matter as the claims fail to assert a practical application to the invention. Further, claims 15-18 are directed to a signal bearing medium that clearly includes carrier signal which is non-statutory subject matter.”

Applicants respond to the Examiner’s above analysis, as follows.

First, Applicants respectfully submit that the above-recited rejection contains a fundamental underlying flaw, as follows: the rejection above is based upon the rejection of the Examiner’s characterization of the claimed invention, not the plain meaning of the claim language itself. That is, not even the original wording of the independent claims has anything to do with a calculation that merely results in numbers.

Rather, the present invention relates to managing the data for calculation in a manner that addresses the problem of cache thrashing. The Examiner simply re-defines the claimed invention.

Applicants respectfully submit that it is clear that any invention, given a reasonably creative Examiner, could be characterized in a manner so as to place it outside statutory subject matter. Therefore, Applicants respectfully request that the Examiner reconsider the rejection currently of record and/or justify on the record, prior to proceeding to Appeal, how the Examiner’s characterization properly reflects the plain meaning of the language of the claimed invention, even in the original claim wording.

Second, relative to claims 15-18, Applicants submit that these claims are directed to a “...signal-bearing medium tangibly embodying a program of machine-readable instructions executable by a digital processing apparatus” Such claims are usually considered to be “Beauregard claims”, after *In re Beauregard*, 53 F.3d 1583 (Fed. Cir.1995).

In that case, the Commissioner of the USPTO submitted to the Court that Beauregard’s appeal to the Federal Circuit should be dismissed because the USPTO had changed its mind on the issue of the claimed invention. That is, the Commissioner stated to the court “... *that computer programs embodied in a tangible medium, such as floppy diskettes, are patentable subject matter under 35 U.S.C. §101 and must be examined under 35 U.S.C. §§ 102 and 103.*”

Thus, the Commissioner, in effect, reversed his Board’s decision, submitting to the Court

that the USPTO no longer supported a policy that was the basis for the Appeal. US Patent 5,710,578 was issued on January 20, 1998, to Beauregard, et al. The Examiner might want to read this case holding (since it is only three paragraphs in length), as well as review the patent that issued to Beauregard.

Relative to the Examiner's allegation that claims 15-18 in the present application "... *clearly includes carrier signal which is non-statutory subject matter*", Applicants respectfully submit that the plain meaning of the language of the preamble of claim 15 would not seem to support the Examiner's allegation. That is, carrier signals are not commonly considered by one having ordinary skill in the art as being tangibly embodied in a signal-bearing medium.

Accordingly, Applicants again respectfully traverse this rejection for claims 15-18.

Third, relative to the Examiner's rationale that the claims of the present invention recite a final result that provides merely numerical values, Applicants respectfully bring to the Examiner's attention that the result in *State Street*, which the Examiner characterizes as "dollar amounts to a final share price", is also "merely numerical values". Indeed, Applicants submit that all computers inherently provide a result that is "merely numerical values", since computers are based upon the binary number system and results will always be binary numbers.

Applicants further submit that what the Court was attempting to convey in the *State Street* holding was that the "numerical values" had "real world" application, rather than being abstract ideas. Applicants submit that such "real world application" is clearly present in the present invention.

That is, Applicants submit that even the original claim language of the present application was not directed to processing a mathematical algorithm in the abstract or of providing "merely numerical values." Rather, the present invention is directed to improving efficiency of processing by arranging data in a prescribed manner that reduces thrashing of data in the cache.

This result of improving efficiency in computer processing is clearly a "real world" result, since it can actually be measured. As co-inventor Gustavson explained in his presentation (e.g., on page 4 of the interview presentation for this application) during the aforementioned telephone interview dated February 1, 2007, the present invention provided 92% peak, in comparison to the 75% peak that resulted from normal LAPACK processing.

Moreover, the useful, concrete, and tangible nature of the present invention is clearly described in even the independent claims, since the present invention increases efficiency of the

calculations for linear algebra subroutines by storing the operands in memory in blocks of data based on the dimensions of the cache, preferably as contiguous data in memory and preferably retrieved from memory in increments of line size, such that the data in each contiguous block of data is retrieved as a unit of data. The invention is clearly non-obvious over the cited prior art, since the block size is based upon one dimension that exceeds the dimension of the cache.

Moreover, the real world aspect of linear algebra processing on computers is clearly discussed at lines 9-12 of page 3 of the disclosure, as having uses throughout scientific and engineering fields and even into games and graphics rendering. The present invention increases the efficiency of the matrix processing by its technique of storing data in memory of the computer as blocks of contiguous matrix data to be retrieved as increments of the block data. The present inventors have recognized that matrix processing is inefficient in Fortran and C, as the matrix blocks, as actually stored in memory for Fortran and C, are typically not contiguous in memory, and the matrix data is likewise normally not stored contiguously in memory. Also, the processing of the matrix data will ultimately need to know how the matrix data is laid out in memory.

As defined by the independent claims, the present invention is actually addressed to memory management of data used in such linear algebra processing and is not at all an attempt to preempt a mathematical algorithm in the abstract. It addresses the practical problem of cache data thrashing, as observed by the present inventors, when conventional standard subroutines are processing matrix data. Thus, another way to view the usefulness and tangible nature of the present invention is its result of providing improved efficiency in the processing of matrices on a computer using such standard matrix subroutines. Contrary to the characterization of the Examiner in paragraph 2 on page 2 of the Office Action, the present, therefore, does indeed provide a “real world result”, by it solution to data thrashing during computer execution of standard linear algebra subroutines.

In view of the foregoing, the Examiner is respectfully requested to reconsider and withdraw these rejections for non-statutory subject matter.

III. THE PRIOR ART REJECTIONS

The Examiner alleges that Myszewski teaches the claimed invention described by claims 1-3, 6, 7, 9, 10, 12, 14, 15, and 18, and renders obvious the invention described by claims 5, 8, 13, 16, 19, 21, and 22. Applicant submits, however, that there are elements of the claimed invention which are neither taught nor suggested by Myszewski and that the rejection currently of record fails to meet the initial burden of a *prima facie* rejection.

More specifically, the Examiner points to lines 55-60 of column 4. However, Applicants submit that this description states executing matrix data in standard format in units of block data, possibly using block size related to the size of the cache. However, Applicants submit that this description does not satisfy the plain meaning of the independent claims, since there is no suggestion of actually storing the data in memory in these blocks of data in the form of the blocks to be used in the processing. As pointed out above, the execution of standard matrix subroutines will cause thrashing due to inefficiency caused by storing the matrix data in standard format.

Indeed, the invention of Myszewski is itself constrained to prepare data for the standard DGEMM interface, which uses the Fortran and C data layouts, and has the problems discussed earlier with data thrashing.

The present invention teaches the technique of composite blocking, as articulated in newly added claim 23. The “double blocking” defined in claims 4, 11, 17, and 20 is a specific embodiment of composite blocking. There is no suggestion in Myszewski to convert and store in memory the matrix data of a matrix in square blocks larger than cache and consisting of rectangular blocks of contiguous data that will each fit into the working area of a cache (even though one dimension of the rectangular blocks is larger than its corresponding cache dimension).

To consider this prior art reference in a different perspective, a rectangular block makes DGEMM run faster. However, the DLFA requires square blocking. The composite blocking of the present invention serves both purposes. Myszewski does not even recognize this disparate requirement and does not suggest a technique that satisfies this dual requirement, nor does it address DLFA.

The technique of the present invention reduces the data thrashing that typically occurs when matrix data is simply retrieved in its original format as is required by higher level

languages such as Fortran and C.

The data thrashing occurs, in the conventional processing, because portions of the matrix data are typically flushed from cache during processing so that additional memory accesses are required as the flushed data becomes needed for current processing and because matrix data is typically not contiguous as actually stored in memory. The present invention addresses this data thrashing by actually storing in memory the matrix data as rectangular contiguous sub blocks of a size that fits into cache and, if the matrix data is also contiguous, it will also be organized on memory line size, thereby increasing the speed at which future matrix data will enter the cache to be consumed in the matrix processing. This occurs because the data actually brought into cache from memory will be completely contiguous matrix data (rather than standard layout data in portions of some lines of data) and because the matrix subroutine processing will fit the processing result back into the data currently being consumed, rather than inadvertently flushing out matrix data not yet being processed.

In contrast to the conventional method of dealing with matrix data, the present invention teaches storing the matrix data in memory as actually being rectangular sub blocks of data that fits the sub blocks together so that all the matrix data is contiguous, where "contiguous" means that the data is in the order to be used in its processing. The standard format of higher level conventional computer languages makes no attempt to place matrix data into its appropriate "contiguous" format.

Moreover, Applicants submit that the rejection currently of record also fails to point to specific lines and columns in Myszewski that satisfy the plain meaning of the language of any of the rejected dependent claims, and respectfully requests that the Examiner provide specific reference to line/columns prior to proceeding to Appeal. The rejection currently of record simply makes conclusory statements that are not supported in the reference, since the present invention is actually dealing with inefficiencies caused by storing data as required by higher level language implementation of matrix processing, as typically done in Fortran or C, where data must be stored in standard format.

The Myszewski patent describes standard cache blocking. The ideas behind standard cache blocking were first discovered and reduced to practice during 1984 and 1985 by the IBM Engineering and Scientific Subroutine Library (ESSL) programming team. In February 1986 IBM released the ESSL programming library to the public. The designers of ESSL recognized the

need for level 3 BLAS and helped to propose as an Industry Standard the Level 3 BLAS. ESSL had a version of DGEMM even before the standard was adopted.

Myzewski does not mention any format changes to DGEMM's operands, and analysis of his codes proves that he is using the formats required by DGEMM. It is a well known fact that certain DGEMM operands will thrash any L1 cache. By using the data structure of the present invention, one can prove that it minimizes L1, L2, and TLB misses.

Therefore, in its exemplary embodiment, the present invention is directed at improving the performance of DLFA. In the high performance DLA community, BLAS 3 were invented to make DLFA run faster. To avoid the performance flaw described above, one can use the new data structure (e.g., memory storage technique) described in the present invention. Composite blocking, the subject of the present invention, is a means to increase DLFA performance when using Square Blocks (SB). Note that a SB is a fundamental building block of the data structure. An innovation in the present invention is to realize that a SB can also be viewed as a collection of rectangular contiguous blocks, each fitting into the working area of L1 cache (typically, the major area of cache), even though the rectangular block has one dimension that is larger than the corresponding cache dimension. Also, rectangular blocks are preferred by DGEMM kernels, as its function $C = C - AB$ has an asymmetry.

Moreover, one exemplary key feature of the present invention is that input operands of all DGEMM calls be stored as contiguous data in memory. A major use of the present invention is for producing high performance Dense Linear Algebra Factorization Algorithms (DLFA), since the standard full format data structures of Dense Linear Algebra (DLA) hurt the performance of its factorization algorithms. Full format rectangular matrices are the input and output of the level 3 BLAS.

It follows, therefore, that the LAPACK and Level 3 BLAS approach has a basic performance flaw. By using composite blocking described in the present invention, this flaw can be removed.

In contrast, Myszewski was directed towards standard cache blocking, since the inputs to BLAS3 DGEMM are standard full format arrays.

As mentioned earlier, measurable improvement in efficiency has been measured by using the present invention for Cholesky factorization as an example of a DLFA, wherein, on the IBM Power3 processor Cholesky factorization achieves 92% of peak performance, whereas

conventional full format LAPACK DPOTRF achieves 77% of peak performance.

Hence, turning to the clear language of the claims, in Myszewski there is no teaching or suggestion of: “...storing data contiguously for a matrix subroutine call in a computer memory in an increment block size that is based on a cache size of said computer, a first dimension of said block being larger than a corresponding first dimension of said cache and a second dimension of said block being smaller than a corresponding second dimension of said cache, such that said block fits into a working space of said cache”, as required by independent claim 1. The remaining independent claims have similar language.

Therefore, Applicants submit that there are elements of the claimed invention that are not taught or suggest by Myszewski, and the Examiner is respectfully requested to withdraw these rejections.

IV. FORMAL MATTERS AND CONCLUSION

In view of the foregoing, Applicant submits that claims 1, 2, and 4-27, all the claims presently pending in the application, are patentably distinct over the prior art of record and are in condition for allowance. The Examiner is respectfully requested to pass the above application to issue at the earliest possible time.

Should the Examiner find the application to be other than in condition for allowance, the Examiner is requested to contact the undersigned at the local telephone number listed below to discuss any other changes deemed necessary in a telephonic or personal interview.

The Commissioner is hereby authorized to charge any deficiency in fees or to credit any overpayment in fees to Assignee's Deposit Account No. 50-0510.

Respectfully Submitted,



Date: February 3, 2007

Frederick E. Cooperrider
Registration No. 36,769

McGinn Intellectual Property Law Group, PLLC
8321 Old Courthouse Road, Suite 200
Vienna, VA 22182-3817
(703) 761-4100
Customer No. 21254

CERTIFICATION OF TRANSMISSION

I certify that I transmitted electronically, via EFS, this Amendment under 37 CFR §1.116 to Examiner C. Ngo on February 3, 2007.



Frederick E. Cooperrider
Registration No. 36,769